

13955053

Basic Patent (No,Kind,Date): JP 9223192 A2 970826 <No. of Patents: 002>

Patent Family:

Patent No	Kind	Date	Applic No	Kind	Date
JP 9223192	A2	970826	JP 9626952	A	960214 (BASIC)
JP 2856702	B2	990210	JP 9626952	A	960214

Priority Data (No,Kind,Date):

JP 9626952 A 960214

PATENT FAMILY:

JAPAN (JP)

Patent (No,Kind,Date): JP 9223192 A2 970826

PATTERN RECOGNITION DEVICE (English)

Patent Assignee: GIJUTSU KENKYU KUMIAI SHINJOHO; TOKYO SHIBAURA ELECTRIC CO

Author (Inventor): NATORI NAOKI; MIZUTANI HIROYUKI

Priority (No,Kind,Date): JP 9626952 A 960214

Applic (No,Kind,Date): JP 9626952 A 960214

IPC: * G06K-009/62; G06F-015/18; G06T-007/00

Derwent WPI Acc No: * G 97-476947; G 97-476947

Language of Document: Japanese

Patent (No,Kind,Date): JP 2856702 B2 990210

Patent Assignee: GIJUTSU KENKYU KUMIAI SHINJOHO; TOKYO SHIBAURA ELECTRIC CO

Author (Inventor): NATORI NAOKI; MIZUTANI HIROYUKI

Priority (No,Kind,Date): JP 9626952 A 960214

Applic (No,Kind,Date): JP 9626952 A 960214

IPC: * G06K-009/62; G06F-015/18; G06K-009/66

Language of Document: Japanese

THIS PAGE BLANK (USPTO)

DIALOG(R) File 347:JAPIO
(c) 2001 JPO & JAPIO. All rts. reserv.

05608392 **Image available**

PATTERN RECOGNITION DEVICE

PUB. NO.: 09-223192 [J P 9223192 A]
PUBLISHED: August 26, 1997 (19970826)
INVENTOR(s): NATORI NAOKI
MIZUTANI HIROYUKI
APPLICANT(s): GIJUTSU KENKYU KUMIAI SHINJOHO SHIYORI KAIHATSU KIKO [000000]
(A Japanese Company or Corporation), JP (Japan)
TOSHIBA CORP [000307] (A Japanese Company or Corporation), JP
(Japan)
APPL. NO.: 08-026952 [JP 9626952]
FILED: February 14, 1996 (19960214)
INTL CLASS: [6] G06K-009/62; G06F-015/18; G06T-007/00
JAPIO CLASS: 45.3 (INFORMATION PROCESSING -- Input Output Units); 45.4
(INFORMATION PROCESSING -- Computer Applications); 45.9
(INFORMATION PROCESSING -- Other)

ABSTRACT

PROBLEM TO BE SOLVED: To enable more accurate recognition and improve the reliability of decision making by giving an indication to an intermediate part so that the recognition is performed in consideration of the degree of a gaze at each partial area.

SOLUTION: The intermediate part 2 learns a partial vector $I_{(sub i)}$ supplied from an input part 1 in learning mode, and recognizes the partial vector $I_{(sub i)}$ supplied from the input part in recognition mode according to the said learning result and outputs the recognition result $R_{(sub i)}$. Then an output part 3 receives the recognition result $R_{(sub i)}$ supplied from the intermediate part 2 and outputs an output signal $O_{(sub j)}$ ($1 < j < m$, where (m) is the number of character categories) corresponding to the recognition result $R_{(sub i)}$. Further, a feedback part 4 determines the 'degree of a gaze' at each partial area of a pattern by feedback trial unless the output signal $O_{(sub j)}$ received from the output part 3 is clear, and stores or learns it to give the 'degree of a gaze' at each partial area at the time of the recognition of a subsequent pattern and instruct rerecognition.

THIS PAGE BLANK (USPTO)

(10)日本国特許庁 (J P)

(12) 特許公報 (B 2)

(11)特許番号

第2856702号

(45)発行日 平成11年(1999)2月10日

(24)登録日 平成10年(1998)11月27日

(51)Int.Cl.
G 0 6 K 9/62
G 0 6 F 15/18
G 0 6 K 9/68

識別記号
6 2 0
5 6 0

F 1
G 0 6 K 9/62
G 0 6 F 15/18
G 0 6 K 9/66

6 2 0 D
5 6 0 A

請求項の数3(全15頁)

(21)出願番号 特願平8-26952

(22)出願日 平成8年(1996)2月14日

(65)公開番号 特開平9-223192

(43)公開日 平成9年(1997)8月26日
審査請求日 平成8年(1996)2月14日

(73)特許権者 593162453

技術研究組合新情報処理開発機構
東京都千代田区東神田2-5-12 鎮角
館ビル8階

(73)特許権者 000003078

株式会社東芝
神奈川県川崎市幸区堀川町72番地

(72)発明者 名取 直毅

神奈川県川崎市幸区柳町70番地 株式会
社東芝柳町工場内

(72)発明者 水谷 博之

神奈川県川崎市幸区柳町70番地 株式会
社東芝柳町工場内

(74)代理人 弁理士 鈴江 武彦

審査官 酒井 伸芳

最終頁に続く

(54)【発明の名称】 パターン認識装置

1

(57)【特許請求の範囲】

【請求項1】 入力部より与えられたパターンについて認識を行う中間部と、該中間部によって認識された認識結果に基づいて出力信号を出力する出力部とを有するパターン認識装置において、

前記出力部の出力信号からパターン特定ができないときに、前記パターンの各部分領域それぞれに対する注視度を選択する注視判断部と、

前記注視判断部が選択した前記各部分領域に対する注視度を加味した認識をさせるべく前記中間部に指示する注視度制御部と、を具備することを特徴とするパターン認識装置。

【請求項2】 入力部より与えられたパターンについて認識を行う中間部と、該中間部によって認識された認識結果に基づいてパターンの種別毎の可能性の高さを示す

2

出力信号を出力する出力部とを有し、当該出力信号値よりパターン判定を行うようにしたパターン認識装置において、

前記出力部から得られる出力信号が二種以上のパターン種別のいずれでもある可能性を示唆するときに、前記パターンを複数の部分領域に分けたときの各部分領域それぞれに対する注視度を選択する注視判断部と、

前記注視判断部が選択した前記各部分領域に対する注視度を加味した認識をさせるべく前記中間部に指示する注視度制御部と、を具備することを特徴とするパターン認識装置。

【請求項3】 入力部より与えられたパターンについて認識を行う中間部と、該中間部によって認識された認識結果に基づいてパターンの種別毎の可能性の高さを示す出力信号を出力する出力部とを有し、当該出力信号値よ

THIS PAGE BLANK (USPTO)

りパターン判定を行うようにしたパターン認識装置において、前記出力部から得られる出力信号が二種以上のパターン種別のいずれでもある可能性を示唆するときに、前記パターンを複数の部分領域に分けたときの各部分領域それぞれに対する注視度を選択する注視判断部と、前記注視判断部より与えられた各部分領域に対する注視度を加味した認識をさせるべく前記中間部に指示する注視度制御部と、前記出力部の出力信号を複数保存し、この保存した出力信号からパターン種別の判断結果を出力する出力信号累積計算部と、を具備することを特徴とするパターン認識装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、文字パターンや画像パターン（以下「パターン」という）を認識するパターン認識装置に関する。

【0002】

【従来の技術】従来のパターン認識の手法の一つとして、パターンをベクトルと看做し、ベクトル空間内での分布領域を学習することにより、該入力パターンを認識するようにしたものがある。

【0003】これについて、図6を参照して説明する。ここでは文字認識を例に挙げて説明する。文字のパターンは、図6に示すようにドット・パターンで示すことができる。この図6の文字パターンは‘2’なる数字のパターン例を示しているが、この例のようにドット・パターンにおいて、白い部分のドットは‘0’、黒い部分のドットは‘1’とし、ドットを1次元に展開すれば、文字パターンは例えば次に示すような‘0’と‘1’からなるベクトルとなる。

【0004】[0, 0, 1, 1, 1, 1, 0, 0, · · ·]

今、このような文字パターンのベクトルの任意のものをXとすると、パターン認識装置は、入力された文字パターンがベクトル空間内のどのような領域にあるかで、この入力文字を分類する。特定の文字種がベクトル空間内のどの領域を占めるかは、一般に現場から収集した多数の文字データを用いて学習する。

【0005】図7は、このようなベクトル空間を説明するための概念図である。同図において、パターン空間を2次元と仮定しており、入力文字パターンベクトルXが、‘4’と分類されている場合を示している。

【0006】図8は、近年注目されているニューラルネットによる文字認識方式の原理を説明するための図であり、これも前述の文字認識と同一原理である。図8は数字認識用のニューラルネットの一例を示したものであり、入力層ILと、中間層MLと、出力層OLを有する階層型のニューラルネットである。入力層ILの各ユニ

ットには前述の文字パターンベクトルXの各要素の値が入力されるが、この場合、入力層ILのユニット数は入力文字パターンベクトルXの次元数に一致する。

【0007】ニューラルネットでは、多くの文字パターンを学習するが、それは前述のような文字パターンベクトルの空間において、‘0’と見なせるベクトル空間のどの領域を占めるか、‘1’と見なせるベクトル空間のどの領域を占めるかなど、ベクトル空間を‘0’、‘1’、‘2’、·····、‘9’に対応する領域に分割していることになる。

【0008】一般に、そのような領域は、超球と呼ばれる中間層MLのユニットの反応領域を単位として構成される。図9は、ベクトル空間における超球を説明するための図である。同図は、図7と同様に、パターン空間を仮に2次元平面として表したものであり、超球は、円で表されている。

【0009】つまり、ニューラルネットは、これら超球状の反応領域の組み合わせでパターンの分布領域を学習していると言える。従って、出力層OLの各ユニットは、文字種に対応するものであり、この場合には、‘0’、‘1’、‘2’、·····、‘9’に対応する10個のユニットである。

【0010】一般に、各出力ユニットは、“0.0”以上“1.0”以下の実数値を出力し、それが“該文字パターンベクトルXのカテゴリである可能性の高さ”を示している。通常、判定に当たってはこれらのうち、最も高い値を出した出力ユニットを選択する。そして、その出力ユニットの対応する文字種を、文字種の判定結果とする。

【0011】以上述べたニューラルネットによるパターン認識装置は、既に文字認識などに利用されているが、出力結果が明確でないとき、つまり、異なるいくつかの文字種対応の出力ユニットがそれぞれ似通った高い値を示している場合に誤判定をするおそれがある。例えば入力文字として、崩れた手書きの‘2’という数字が与えられ、この文字の場合に、‘2’とも、あるいは‘7’とも読めたとする。

【0012】このような場合には、‘2’なる文字の可能性と、‘7’なる文字の可能性とを有することから、

40 当該崩れた手書きの‘2’の文字パターンベクトルXの入力に対して、文字種‘2’の可能性を示す出力ユニットOL₂と、文字種‘7’の可能性を示す出力ユニットOL₇の出力値がほぼ等しい値となる。そして、たとえ微差であっても、出力ユニットOL₂と出力ユニットOL₇とで差があるときはその高い方の値を示す側を採用することとなる。

【0013】故に、文字種が‘2’であるはずが、‘7’の文字パターンベクトルXの方がわずかの差で‘2’の文字パターンベクトルXより高い値を呈していれば、高い方を判定結果としてしまい、‘2’の文字バ

THIS PAGE BLANK (USPTO)

ターンベクトルXを「'7'である」と誤判定してしまうという問題があった。

【0014】このような場合に、従来では必要なときにフィードバック機構を用いて認識を修正するようにしたものがある。図10は、フィードバック機構を用いた従来のパターン認識装置を示したものである。

【0015】この装置においては、フィードバック機構であるフィードバック部4は、出力部3より与えられた出力信号をもとにフィードバックによる認識の修正が必要かどうかを判断する。そして、フィードバックが必要であれば、中間部2に認識を修正するよう指示する。

【0016】ここで、"認識を修正すること"を"注視"といい、"注視の度合"を"注視度"ということにする。しかし、従来のフィードバック機構を用いたパターン認識装置では、文字パターン全体を複数の領域に分けた部分領域のうち、1つのみを選び、それについての"注視度"を制御するだけであった。つまり、複数の部分領域の中から1つのみを選び、それについての"注視度"を制御するだけであった。

【0017】例えば、ある出力信号の内容から i ($1 \leq i \leq n$; 但し、 n は部分領域数) 番目の部分領域についての認識結果が明確であると判断されたときは、 i 番目の部分領域の注視度を大きくする、などの動作である。

【0018】しかし、"注視"する部分領域からその周辺の部分領域への"注視度"の変化が不連続であり、それが"注視"を不安定にするという問題があった。また、1箇所あるいは1回だけの"注視"が、必ずしも問題に対する最適解を与えるとは限らないという懸念を残すという課題もある。つまり、上記の例で言えば、 i 番目と j ($1 \leq j \leq n$, $j \neq i$) 番目の部分領域の"注視度"を大きくすることが最適な判定をもたらすかもしれないということである。

【0019】

【発明が解決しようとする課題】このように、従来のフィードバック機構を用いたパターン認識装置では、文字パターンの認識に当たり、その文字パターンの種々の部分領域の中から1つのみの部分領域を選び、それについての"注視度"を制御するだけであった。

【0020】しかし、"注視"する部分領域からその周辺の部分領域への"注視度"の変化が不連続であり、それが"注視"を不安定にするという問題があった。また、1箇所あるいは1回だけの"注視"が必ずしも問題に対する最適解を与えるとは限らず、他の部分領域の"注視度"を大きくすることが最適な判定をもたらすかも知れないなど、認識処理にあたり改良の余地を残している。

【0021】そこで、本発明は、より正確な認識を可能とすると共に、信頼性の高い判定を実現することができるようにしたパターン認識装置を提供することを目的と

する。

【0022】

【課題を解決するための手段】上記目的を達成するため、本発明は次のようにする。すなわち、入力部より与えられたパターンについて認識を行う中間部と、該中間部によって認識された認識結果に基づいてパターンの種別毎の可能性の高さを示す出力信号を出力する出力部とを有し、当該出力信号値よりパターン判定を行うようにしたパターン認識装置において、前記出力部から得られる出力信号が二種以上のパターン種別のいずれでもある可能性を示唆するときに、前記パターンを複数の部分領域に分けたときの各部分領域それぞれに対する注視度を選択する注視判断部と、前記注視判断部が選択した前記各部分領域に対する注視度を加味した認識をさせるべく前記中間部に指示する注視度制御部とを具備するようにした。

【0023】また、前記出力部の出力信号を複数保存し、この保存した出力信号からパターン種別の判断結果を出力する出力信号累積計算部を設けるようにした。本発明は、安定した注視ができるよう、パターンの各部分領域に対する注視度をフィードバック試行により定め、それを蓄積または学習することで、これに基づいて以後のパターンの認識の際に各部分領域に対する最適な注視度を与えて認識を行い、また、保存された複数の出力信号を用いて判定を修正し得るようにしたパターン認識装置であり、入力部より与えられたパターンについて中間部にて認識を行うと共に、出力部は該中間部によって認識された認識結果に基づいてパターンの種別毎の可能性の高さを示す出力信号を出し、この出力信号からパターンの判定を行うが、前記出力部の出力信号が明確でないときに、つまり、出力信号が似通っていて複数のパターン種別いずれもが該当する可能性があるときに、前記パターンを複数の部分領域に分けた際の各部分領域に対する注視度をフィードバック試行により定め、それを蓄積または学習し、これに基づいて以後のパターンの認識の際に各部分領域に対する最適な注視度を注視判断部から出力させ、この注視判断部より出力された各部分領域に対する注視度に加味した認識をさせるべく、前記中間部に指示し、また、出力部の出力信号を複数保存して、この保存した出力信号から最終的なパターン種別判断するようにした。

【0024】更には、本発明は、前記入力部は、ニューラルネットワークの入力層によって構成し、前記中間部は、ニューラルネットワークの中間層によって構成し、前記出力部は、ニューラルネットワークの出力層によって構成すると共に、前記出力部の出力信号が明確でないときに、前記パターンの各部分領域に対する注視度をフィードバック試行により定め、それを蓄積または学習して、これに基づき、以後のパターンの認識の際に各部分領域に対する最適注視度を得、この注視度に従って、各

部分領域に対する認識修正を前記中間部に実施させるようし、前記中間部からの出力信号を複数保存して、パターン種別の最終的な判断を行うようにしたニューラルネットワーク型のパターン認識装置とした。

【0025】従って、本発明によれば、出力信号が明確でないとき、パターンの各部分領域に対する注視度をフィードバック試行により定め、それを蓄積または学習することで、以降のパターンの認識の際に各部分領域に対する注視度を与えて再認識を行うことができるため、より正確な認識が可能となる。また、パターンを複数の各部分領域に分けてこれら各部分領域に対する最適な注視度を与えて認識を行うようにしたことにより、従来のような選ばれた1つの部分領域のみを注視する手法よりも信頼性の高い判定を実現することができる。さらに、出力信号を複数保存し、それを用いることとも、信頼性の高い判定をもたらす。

【0026】また、本発明によれば、ニューラルネットワークによる事例学習機能を有し、出力信号が明確でないとき、パターンの各部分領域に対する注視度をフィードバック試行により定め、それを蓄積または学習することで、以降のパターンの認識の際に各部分領域に対する注視度を与えて再認識を行うようにすることができるため、より正確な認識が可能となる。

【0027】また、複数ある各部分領域それぞれに適した注視度を与えるようにしたことにより、従来のような選ばれた1つの部分領域のみを注視する手法よりも信頼性の高い判定を実現することができる。さらに、出力信号を複数保存し、それを用いることとも、信頼性の高い判定をもたらす。

【0028】さらに、本発明によれば、注視判断部をニューラルネットワークによって構成するため、ニューラルネットワークの特徴である汎化性や耐ノイズ性を持つ注視を得ることができる。このとき、パターンの各部分領域に対する注視度をユーザが明示的に修正することができるので、この修正によりユーザの意図を注視に反映させることができるようにになる。また、パターンの各部分領域に対する注視度を検証し、自動的に修正するようにすることもできるので、コンフリクトのない最適な注視を実現することができる。

【0029】

【発明の実施の形態】以下、図面を参照して本発明の具体例について説明する。なお、図を簡略化するため、ベクトル、信号など複数の要素からなるデータの伝達路であっても1本の線で表現することができる。ここでは文字認識を例に挙げて説明する。

【0030】図1は、本発明の具体例に係るパターン認識装置の概略構成を示す図であり、入力部1、中間部2、出力部3、フィードバック部4から構成されている。これらのうち、入力部1は、文字パターンベクトルXを入力すると、これを複数の部分領域に分け、それぞ

れの部分領域の文字パターンベクトルである部分ベクトル I_i ($1 \leq i \leq n$; nは部分領域数)を出力するものである。文字パターンベクトルXは、図6に示すようにドット・パターンで示すことができる。このドット・パターンにおいて、白い部分のドットは'0'、黒い部分のドットは'1'とし、ドットを1次元に展開すれば、文字パターンベクトルXは例えば次に示すような'1'と'0'からなるベクトルとなる。

【0031】[0, 0, 1, 1, 1, 1, 0, ..., ...]

ここで、学習モード時に入力される文字パターンベクトルXには、教示信号Tが付加されて入力部1に入力される。教示信号Tは、文字パターンベクトルXが何の文字種であるかを装置に教示するための信号であり、例えば、数字を学習する装置に、そのパターンは'2'である、と教示するには、「[2]」といった如きの教示信号Tを用意し、文字パターンベクトルXの先頭あるいは末尾などにこの教示信号Tを付加する。

【0032】つまり、先の例の場合、

20 [0, 0, 1, 1, 1, 1, 0, ..., ...] [2] あるいは

[2] [0, 0, 1, 1, 1, 1, 0, ..., ...] といった如きである。

【0033】部分ベクトル I_i は、文字パターンベクトルXの要素の一部からなるベクトルである。学習モード時には、各部分ベクトル I_i の先頭、あるいは末尾などに教示信号Tが付加される。

【0034】中間部2は、学習モード時には入力部1から与えられる部分ベクトル I_i を学習し、認識モード時にはこの学習結果をもとに、入力部1から与えられる部分ベクトル I_i を認識し、その認識結果 R_i を出力するものである。学習モード時には、各認識結果 R_i の先頭あるいは末尾などに教示信号Tが付加される。

【0035】出力部3は、中間部2から与えられた認識結果 R_i を受けて、その認識結果 R_i 対応の出力信号 O_j ($1 \leq j \leq m$; 但し、mは文字カテゴリ数)を出力するものである。学習モード時には、各出力信号 O_j の先頭あるいは末尾などに教示信号Tが付加される。ここで、認識結果 R_i 対応の出力信号 O_j とは、パターン種別毎の可能性の高さを示す値(評価値)の情報を持つ信号であること意味し、認識結果 R_i に対応した値を持つように変換されて出力される信号である。つまり、数字の文字の認識の場合、数字の'1'であることの可能性の高さ、数字の'2'であることの可能性の高さ、数字の'3'であることの可能性の高さ、...を認識結果 R_i に応じて数値表現したものである。

【0036】フィードバック部4は、出力部3から受けた出力信号 O_j が明確でないときに、パターンの各部分領域に対する“注視度”をフィードバック試行によって定め、それを蓄積または学習することで、以降のパター

ンの認識の際に各部分領域に対する“注視度”を与えて再認識を行わせるものである。フィードバック部4は、注視判断部41、注視度制御部42、出力信号累積計算部43を備える。

【0037】なお、上述の“明確でない”とは、例えば、文字認識の場合に、ある特定の文字である可能性の高さを示す数値が、他の別の文字である可能性の高さを示す数値と似たりよったりで、いずれの文字であるかを断定できない状態をいう。つまり、二種以上のパターン種別のいずれでもある可能性がある状態を指す。

【0038】これらのうち、注視判断部41は、出力部3からの出力信号O_jをもとにフィードバックによる注視が必要かどうかを判断すると共に、注視が必要である場合には注視パターンベクトルZを出力して注視度制御部42に与えるものである。

【0039】ここで、注視パターンベクトルZとは、パターンの各部分領域に対する注視度からなるベクトルであり、種々の注視度の値を並べたベクトルである。注視パターンベクトルZは注視度の値やその並べ方を変えた種々のパターンのものを予め用意してあり、注視判断部41は必要に応じてこれらを選択して出力するようにしてある。

【0040】注視度制御部42は、注視パターンベクトルZに従って認識を修正するよう、中間部2に指示したりする等の制御を行うものである。また、出力信号累積計算部43は、フィードバックを複数回、異なる注視パターンベクトルZに従って行われることにより、得られた各出力信号O_jを記憶し、これら各出力信号O_jをもとに最終的な判定信号Yを出力するものである。

【0041】以上のような構成の本装置において、文字パターンベクトルXが入力部1に入力されると、入力部1はこれを部分ベクトルI_iとして中間部2に出力する。そして、中間部2では、学習モード時には部分ベクトルI_iを学習し、認識モード時にはこの部分ベクトルI_iについて何の部分ベクトルであるかを認識してその認識結果R_iを出力する。

【0042】この中間部2から出力された認識結果R_iは出力部3に入力され、出力部3では、この入力された認識結果R_iに応じた出力信号O_jを出力する。出力信号O_jは、学習モード時と認識モード時とで異なる。

【0043】[学習モード時での動作] 学習モード時には出力部3は、まず、出力信号O_jをフィードバック部4の注視判断部41と出力信号累積計算部43に与える。注視判断部41は、まず、出力信号O_jをもとにフィードバックによる注視が必要かどうかを判断する。これは出力信号O_jから、文字パターンの文字種が明確に判定できないと判断されるか否かでフィードバックによる注視が必要か否かを判断する。

【0044】注視判断部41での判断の結果、フィードバックが必要であれば、この注視判断部41は注視度制

御部42に複数回、異なる注視パターンベクトルZを与える。

【0045】注視パターンベクトルZは、パターンの各部分領域に対する注視度からなるベクトルである。例えば、最も低い注視度を“0.0”、最も高い注視度を“1.0”とすれば、ある注視パターンベクトルZは以下のようになる。

[0.0, 0.5, 1.0, 0.5, 0.0, ·····]

この注視パターンベクトルZは、ある位置を最大の注視度に定め、その隣りを中間の注視度に定め、それ以外を最低の注視度に定めるというものである（このような注視パターンベクトルZを種々予め用意して利用する）。

【0046】すると、注視度制御部42は、この与えられた注視パターンベクトルZに従って認識を修正するよう、中間部2に指示する。この指示は、注視パターンベクトルZ_Nを出力して中間部2に与えることで行う。

【0047】注視パターンベクトルZ_Nは、例えば正規化などにより注視パターンベクトルZを実際の制御量に変換したものであるが、注視パターンベクトルZと同値であることもある。

【0048】そして、この指示により中間部2では、注視パターンベクトルZ_Nの示す各部分の注視度を加味して部分ベクトルI_iについての再認識を行う。そして、中間部2での再認識の結果（認識結果R_i）は出力部3に送られ、この出力部3にてこの再認識による認識結果R_i対応の出力信号O_jに変換され、再び注視判断部41と出力信号累積計算部43に与えられることになる。

【0049】そして、出力信号累積計算部43は出力部3からの出力信号O_jを累積計算し、また、注視判断部41は、この出力信号O_jをもとにフィードバックによる注視が必要かどうかを判断する。その判断の結果、フィードバックが必要であれば、注視判断部41は注視度制御部42に前回と異なる注視パターンベクトルZを与える。注視度制御部42は、この与えられた注視パターンベクトルZに従って認識を修正するよう、中間部2に指示する。そして、この指示により中間部2では、注視パターンベクトルZ（実際はZ_N）を加味して部分ベクトルI_iについての再認識を行う。

【0050】そして、中間部2での再認識の結果（認識結果R_i）は出力部3に送られ、この出力部3にてこの再認識による認識結果R_i対応の出力信号O_jに変換され、再び注視判断部41と出力信号累積計算部43に与えられることになる。出力信号O_jを受けた出力信号累積計算部43は、この出力信号O_jを累積計算する。

【0051】一方、出力信号O_jを受けた注視判断部41は、当該出力信号O_jをもとに再認識の有無を判断し、再認識の必要がないと注視判断部41が判断したときは、注視パターンベクトルZに従った認識修正を終了する。

【0052】このようにして、不明確さが解消されるま

11

で、フィードバックを複数回、異なる注視パターンベクトル Z に従って行うことで再認識を繰り返し、それぞれにおいて得られた各出力信号 O_j を出力信号累積計算部43に累積計算して保存するが、この操作をフィードバック試行と呼ぶ。

【0053】注視判断部41は、このようなフィードバック試行により、良好な注視パターンベクトル Z を定め、それを蓄積または学習する。以上が学習モードでの動作である。

【0054】【認識モード時での動作】一方、認識モード時には、出力部3からの出力信号 O_j が注視判断部41と出力信号累積計算部43に与えられると、注視判断部41は、まず、出力信号 O_j をもとにフィードバックによる注視が必要かどうかを判断する。フィードバックが必要であれば、蓄積または学習によって得られた良好な注視パターンベクトル Z を、複数出力する。ここで、出力される注視パターンベクトル Z の個数を“注視回数”という。

【0055】注視度制御部42は、注視パターンベクトル Z に従って、中間部2に認識を修正するよう指示する。この指示は、注視パターンベクトル Z_N を出力することで行う。注視パターンベクトル Z_N は、例えば正規化などにより、注視パターンベクトル Z を実際の制御量に変換したものであるが、注視パターンベクトル Z と同値であることもある。

【0056】そして、中間部2からのこの認識修正指示により、中間部2では注視パターンベクトル Z_N を用いた再認識が行われ、その再認識結果は出力部3を通して出力信号 O_j に変換されて出力される。そして、この出力信号 O_j が注視判断部41と出力信号累積計算部43に与えられる。このようなフィードバックを注視回数行い、各出力信号 O_j が出力信号累積計算部43に保存される。

【0057】出力信号累積計算部43は、各出力信号 O_j をもとに最終的な判定結果である判定信号 Y を出力する。例えば、判定信号 Y は、各出力信号 O_j の積算を行って求めるなどである。

【0058】この結果、例えば認識対象文字のカテゴリが‘2’と‘7’両方について可能性が高いような明確でないものであっても、認識対象文字のパターンの複数の部分領域について、その各部分領域に対する最適な注視度を、フィードバック試行により定め、それを蓄積または学習することで、以降のパターン認識の際に、認識対象文字のパターンの各部分領域に対する最適な“注視度”を与えて再認識を行うことができるようになるため、より正確な認識が可能となる。

【0059】また、認識対象文字のパターンを複数に領域分割し、その各部分領域に対する“注視度”を与えて注視度対応に認識結果を得て、これら各部分領域の認識結果の総合評価から文字種の判定をして判定信号 Y を得

12

るようとしたことにより、従来の選ばれた1つの部分領域を注視する手法よりも信頼性の高い判定を実現することができる。さらに、認識対象文字のパターンについて、その各部分領域の認識結果を複数保存し、それを用いて総合的に判断することも、信頼性の高い判定をもたらす。

【0060】また、注視判断部41の機能はニューラルネットワークを用いて構成することもでき、この場合、ニューラルネットワークの特徴である汎用性や耐ノイズ性を持つ“注視”を得ることができるようになる。このとき、パターンの各部分領域に対する注視度をユーザが明示的に修正することができる、ユーザの意図を注視に反映させることができる。また、パターンの各部分領域に対する注視度を検証し、自動的に修正することができる、コンフリクトのない最適な注視を実現することができる。

【0061】図2は、本発明のより具体的な構成例を示すブロック図である。図2に示すパターン認識装置は、入力部1、中間部2、出力部3、フィードバック部4のうち、中間部2は部分領域認識モジュール $21_1, 21_2, \dots, 21_n$ にて構成されており、また、出力部3は出力信号生成部 $31_1, 31_2, \dots, 31_n$ より構成されている。

【0062】入力部1は、文字パターンベクトル X を入力し、部分ベクトル I_i ($1 \leq i \leq n$; 但し、 n は部分領域数)を出力する。文字パターンベクトル X から部分ベクトル I_i を出力するには、具体的には次のようにしている。

【0063】図3は、上記部分領域の定義を示す図である。部分領域とは、文字パターン画素について、所定の画素サイズで区切られる領域であり、各部分領域は互いに隣接する他の部分領域について、互いに一部重複領域を持つような領域としてある。図3の例の場合、 14×20 画素構成の1つの文字パターン画素を、1から40までの計40個の部分領域に分けてあり、第1番目の部分領域と、第2番目の部分領域とではかなりの部分で互いに重なりあった区分となっていることがわかる。

【0064】また、部分ベクトル I_i とは、例えば、部分領域の範囲を示す枠の、当該枠内に含まれる各画素の濃淡値を並べたものである。本発明における具体例の場合、部分領域は領域を重複させながら例えば、横方向5箇所、縦方向8箇所設けられるので、縦横合計40個の部分領域となる。すなわち、部分領域数 n は、“40”である。

【0065】中間部2は、各部分ベクトル I_i を学習、認識する部分領域認識モジュール $21_1, 21_2, \dots, 21_n$ によって構成されており、学習モード時には部分ベクトル I_i を学習し、認識モード時には部分ベクトル I_i を認識し、認識結果 R_i を出力する。

【0066】図4は、1つの部分領域認識モジュール2

13

1_i ($i = 1, 2, 3, \dots$) の構成を示す図である。図に示すように、1つの部分領域認識モジュール 21_i は、複数のユニット $21_{i01} \sim 21_{i9k}$ より構成されている。そして、例えば、ユニット 21_{i01} は、部分ベクトル I_i を入力し、以下の式に従って認識結果 R_{i01} を出力する。

【0067】

【数1】

$$R_{i01} = ZN_i * \exp(-D_{i01}/\|W_{i01}\|)$$

【0068】これは、図5に示されるような曲線である。ここで、 D_{i01} とは、部分ベクトル I_i とユニット 21_{i01} の荷重ベクトル W_{i01} とのユークリッド距離、すなわち、

【0069】

【数2】

$$D_{i01} = \sqrt{\sum_d (W_{i01d} - I_{id})^2}$$

である。ただし、 W_{i01d} は荷重ベクトル W_{i01} の第d番目の要素、 I_{id} は入力された部分ベクトル I_i の第d番目の要素である。また、

【0070】

【数3】

$$\|W_{i01}\|$$

は、荷重ベクトル W_{i01} の大きさである。 ZN_i は、部分領域認識モジュール 21_i に与えられる注視度であり、正規化されているものとする。なお、学習モード時には、荷重ベクトル W_{i01} を修正することで、パターンを学習させる。最も単純な学習方法は、荷重ベクトル W_{i01} を、入力された部分ベクトル I_i そのものとすることである。これにより入力が、学習した部分ベクトル I_i に近ければ近いほど、ユニット 21_{i01} は大きな値の認識結果 R_{i01} を出力するようになる。

【0071】出力部3は、中間部2から与えられた認識結果 R_i を受けて、この認識結果 R_i 対応の出力信号 O_j ($1 \leq j \leq m$; 但し、 m は文字カテゴリ数) を出力する。例えば、各出力信号 O_j としては、その文字カテゴリである可能性を “0.0” 以上 “1.0” 以下の信号の強さで表現したものとしてある。

【0072】最も単純な方法は、各部分領域認識モジュール 21_i 毎に、認識結果 R_{ijk} ($1 \leq j \leq m$, $1 \leq k \leq K$; 但し、 K は文字カテゴリあたりのユニット数) の中の最大値をとり、それらの和を求めることがある。すなわち、次のような式により、出力信号 O_j を求める。

【0073】

【数4】

14

$$O_j = \sum_{i=1}^n \max_k R_{ijk}$$

【0074】フィードバック部4は、出力信号 O_j が明確でないときに(つまり、別の他のパターンと似通った認識結果となるような値であったとき)、パターンの各部分領域に対する注視度をフィードバック試行によつて定め、それを蓄積または学習することで、以降のパターンの認識の際に各部分領域に対する注視度を与えて再認識を行わせる。

【0075】学習モード時には、まず、出力信号 O_j が注視判断部41と出力信号累積計算部43に与えられる。注視判断部41は、まず、出力信号 O_j をもとにフィードバックによる注視が必要かどうかを判断する。

【0076】例えば、判断は、「出力信号 O_j の内で最も強い値を示した文字カテゴリが、教示信号Tの示す正解カテゴリと異なる場合は、フィードバックを行う」などの如きである。このような判断の結果、フィードバックが必要であれば、注視度制御部42に複数回、異なる注視パターンベクトルZを与えることになる。

【0077】注視度制御部42は、注視パターンベクトルZにしたがって、中間部2の各部分領域認識モジュール 21_i に認識を修正するよう指示する。この指示は、注視パターンベクトルZNを出力することで行う。

【0078】ここで、注視パターンベクトルZNとは、例えば正規化などにより、注視パターンベクトルZを実際の制御量に変換したものであるが、注視パターンベクトルZと同値であることもある。例えば、正規化による注視パターンベクトルZNは、次の式に従って求められる。

【0079】

【数5】

$$ZN_i = Z_i / \sum_{i=1}^n Z_i$$

【0080】そして、注視パターンベクトルZNに従つて中間部2で再認識が行われ、その認識結果は出力部3により認識結果対応の出力信号 O_j として出力され、この出力信号 O_j が注視判断部41と出力信号累積計算部43に与えられる。

【0081】このようなフィードバックを複数回、異なる注視パターンベクトルZに従つて行い、これによつて得られた各出力信号 O_j を出力累積計算部43に保存する(フィードバック試行)。

【0082】注視判断部41は、このようなフィードバック試行により、良好な注視パターンベクトルZを定め、それを蓄積または学習する。例えば、40個の部分領域に対する良好な注視パターンベクトルZは、以下の如きの40個の値の並びによつて与えられることになる。

50

15

一方、認識モード時には、まず、出力部3からの出力信号O₃が注視判断部41と出力信号累積計算部43に与えられる。注視判断部41は、まず、出力信号O₃をもとにフィードバックによる注視が必要かどうかを判断する。例えば、判断は、「出力信号O₃で1番目に大きい値と2番目に大きい値との差がある閾値以下となる場合は、フィードバックを行う」などといった加きである。

【0084】このような判断の結果、フィードバックが必要であれば、蓄積または学習によって得られた良好な注視パターンベクトルZを、注視回数出力する。注視度制御部42は、注視パターンベクトルZに従って、中間部2の各部分領域認識モジュール21iに認識を修正するよう指示する。この指示は、注視パターンベクトルZNを出力することで行う。そして、中間部2で再認識が行われ、出力部3を通して再び出力信号O3が注視判断部41と出力信号累積計算部43に与えられる。

【0085】このようなフィードバックを社視回数行われることによって得られたそれぞれの認識結果の各出力信号 O_j が、出力信号累積計算部43に保存される。出力信号累積計算部43は、各出力信号 O_j をもとに最終的な判定信号Yを出力する。

【0086】出力信号累積計算部43での判定信号Yの求め方は、フィードバックを行わなかった場合は、最初の出力信号O_jのみを用いて求め、フィードバックを行った場合は、例えば、各出力信号O_jの積算を行って求めるといった具合である。

【0087】以上述べた具体例によれば、入力された認識対象パターンの認識を複数の部分領域にわけて認識し、その各部分領域の認識結果を総合してパターン判別するようにしたものであり、入力された認識対象パターンの認識を行った結果、得られた認識結果対応の出力信号 O_j の値が、明確でない場合、つまり、他の別の種別のパターンと区別のつきにくい似通った値であった場合でも、パターンの各部分領域に対する注視度をフィードバック試行により定め、それを蓄積または学習することで、以降のパターンの認識の際に各部分領域に対する最適な注視度を与えて再認識を行うようにすることができるため、より正確な認識が可能となる。

【0088】また、認識対象パターンを複数の部分領域に分け、各部分領域に対する注視度をそれぞれ与えるようにしてことにより、従来の選ばれた1つのみの部分領域を注視する手法よりも信頼性の高い判定を実現することができる。さらに、認識結果の評価値である出力信号を複数保存し、それを用いて判断することも、信頼性の

16

高い判定をもたらす。

【0089】つぎに、本発明装置における能力について説明する。本発明によるパターン認識装置の能力を検証するために、「0」から「9」までの数字認識実験を行った。まず、代表的な形状の300件の数字文字パターンを用意した。これは、部分領域認識モジュール21iを構成する複数あるユニット21ijkでの荷重ベクトル W_{ijk} を得るためのものである。学習は、「文字パターンベクトル X を荷重ベクトル W_{ijk} そのものとする」、

【0090】すなわち、300件（10種ある数字文字の各カテゴリにつき30件）のパターンを、それぞれ40個の部分領域に分けて学習する。次に、実際の手書き数字パターン36,000件を、フィードバック試行用データとして学習させた。認識用未学習データとして、実際の手書き数字パターン10,000件を用意した。その結果、未学習データ認識率は、次のようなものとなつた。

【0091】すなわち、[従来1(フィードバックなし)]の場合での未学習データ認識率は93.9%、
[従来2(1箇所フィードバック注視)]の場合での未

学習データ認識率は95.6%、そして、[本発明(複数箇所・回数フィードバック注視)]の場合での未学習データ認識率は96.4%である。

【0092】本発明の方式の場合、従来に比べて明らかに、未学習データの認識率向上が図られていることが、確認される。従来のフィードバック機構のない手法では、出力結果（出力信号の値）が明確でないとき、例えば文字種‘2’と‘7’の両方の出力信号値が共に高いといったようなときは、その値に僅かでも差があると、高い値の方を認識結果としてしまい、本当は‘2’のデータであるものを、‘7’であると誤認識されてしまう可能性が高く、安定した認識を実現することは困難となっていた。また、1つの部分領域のみの注視度を制御するフィードバック機構を用いた手法でも、注視の安定性に限界があつた。

【0093】これに対し、本発明によるパターン認識装置では、認識結果として得た出力信号値が他の文字種の値と似通っていて明確でない場合でも、パターンの各部分領域に対する最適な注視度をフィードバック試行により定め、それを蓄積または学習することで、それ以降のパターン認識の際に、各部分領域に対する注視度を与えて再認識を行うようにするため、より正確なパターン認識が可能となった。また、認識対象のパターンを複数の部分領域に分けて、それらの各部分領域に対する注視度をそれぞれ与えて認識を行うようにしたことにより、従来のような選ばれた1つの部分領域を注視する手法よりも、信頼性の高い判定を実現することができた。さら

17

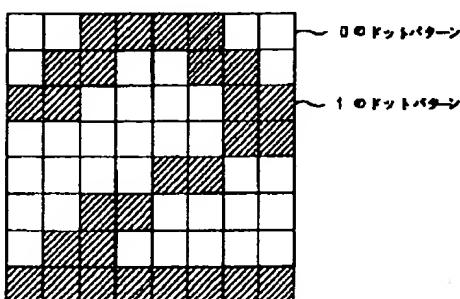
に、フィードバックにより得られたそれぞれの出力信号を複数保存し、それを用いて総合的に判断するようにしたことも、信頼性の高い判定をもたらした。

【0094】なお、本発明は上述した具体例に限定されるものではなく、その要旨を逸脱しない範囲で、種々変形して実施することができる。例えば、中間部2のニューラルネットワークの代わりに、従来の統計的識別装置を用いても良く、また、注視判断部41を構成するニューラルネットワークも、多層型や相互結合型など、さまざまなアーキテクチャを適用して実施することができ

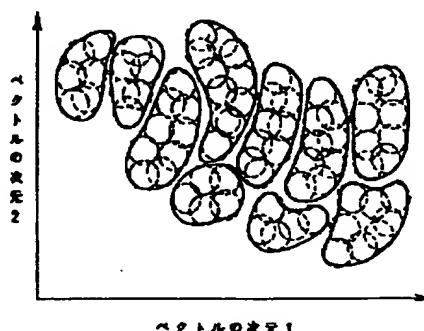
100851

【発明の効果】以上、本発明によれば、出力信号が明確でないときに、パターンの各部分領域に対する注視度をフィードバック試行により定め、それを蓄積または学習することで、以降のパターンの認識の際に各部分領域に対する注視度を与えて再認識を行うため、より正確な認識が可能となる。また、各部分領域に対する注視度を与えることにより、従来の選ばれた1つの部分領域を注視する手法よりも信頼性の高い判定を実現することができる。さらに、出力信号を複数保存しそれを用いて判断することも、信頼性の高い判定をもたらす。また、本発明によれば、ニューラルネットワークによる事例学習機能を有し、出力信号が明確でないときにパターンの各部分領域に対する注視度をフィードバック試行により定め、それを蓄積または学習することで、以降のパターンの認識の際に各部分領域に対する注視度を与えて再認識を行うため、より正確な認識が可能となる。また、各部分領域に対する注視度を与えることにより、従来の選ばれた1つの部分領域を注視する手法よりも信頼性の高い判定を実現することができる。さらに、出力信号を複数保存しそれを用いて判断することも、信頼性の高い判定をもたらす。さらに、本発明によれば、注視判断部をニューラルネットワークによって構成するため、ニューラルネットワークの特徴である汎化性や耐ノイズ性を持つ注視動作を得ることができる。このとき、パターンの各部分

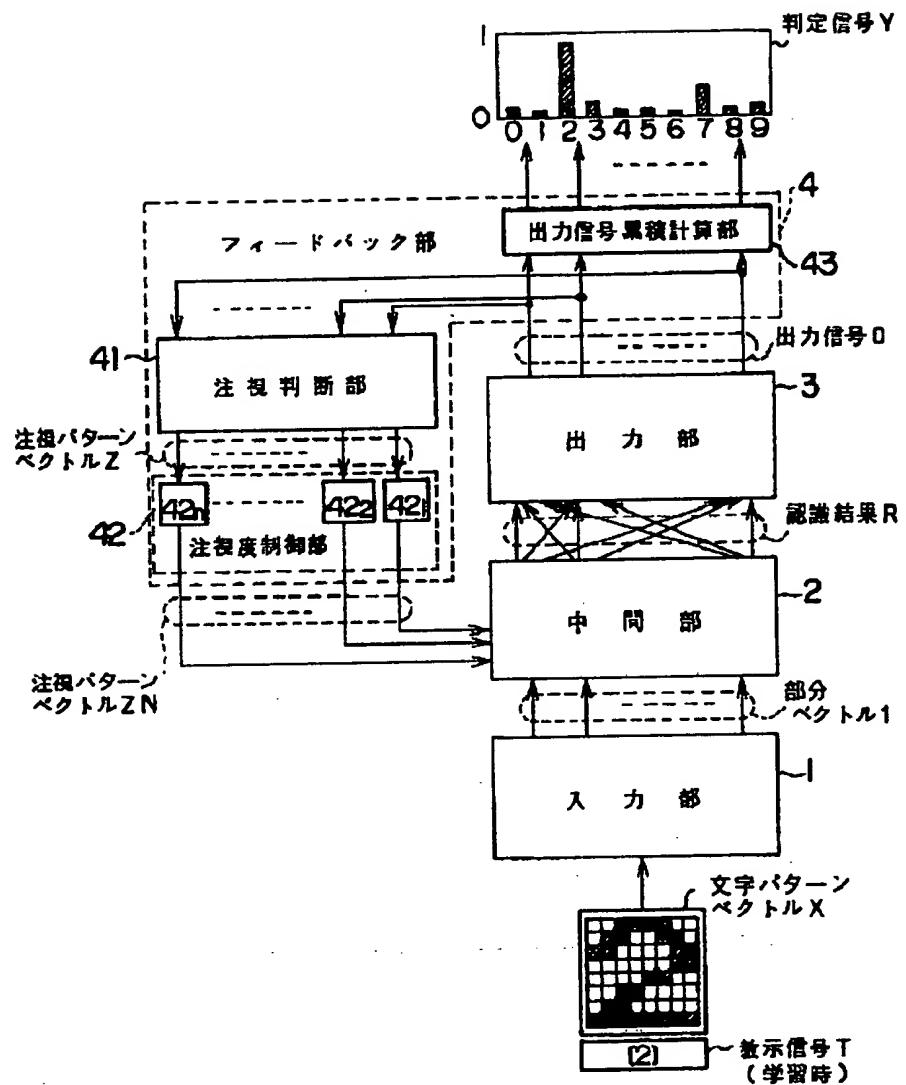
〔圖6〕



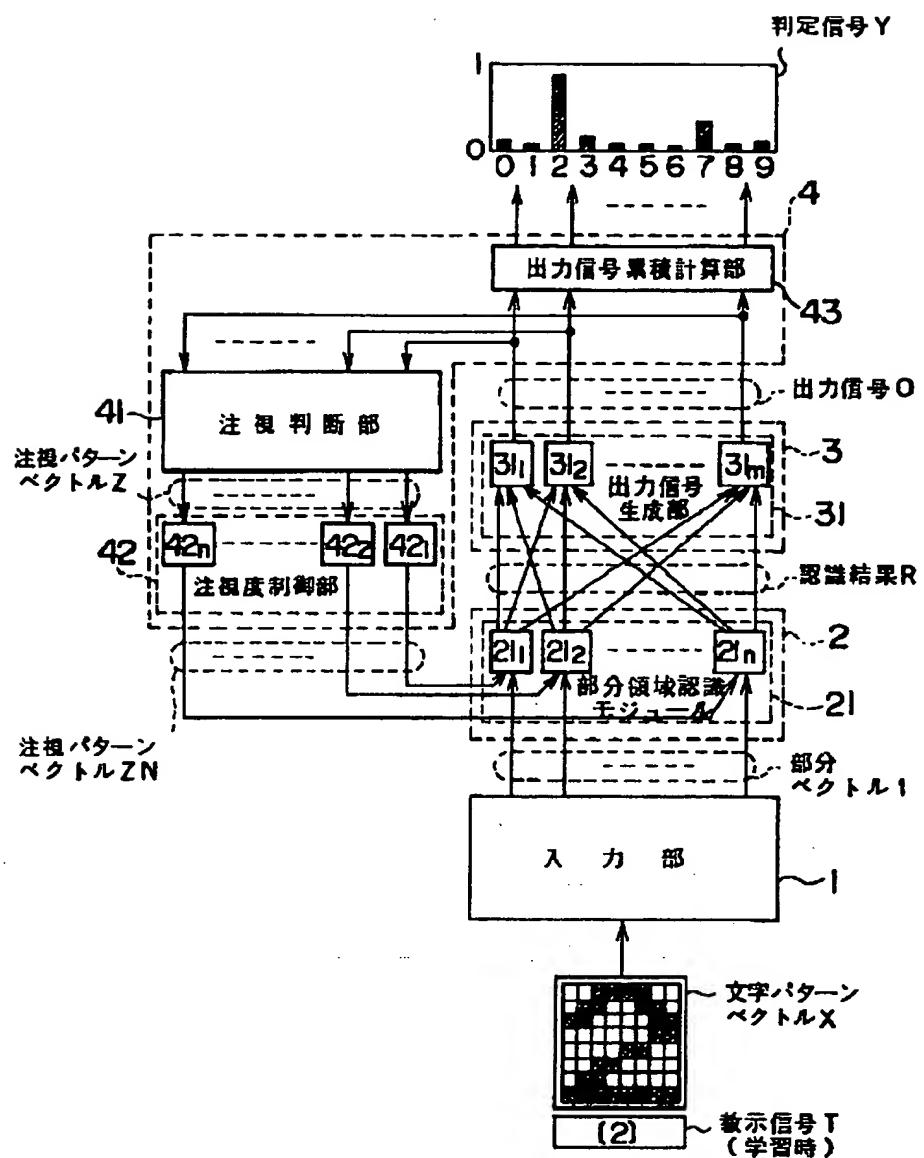
〔四九〕



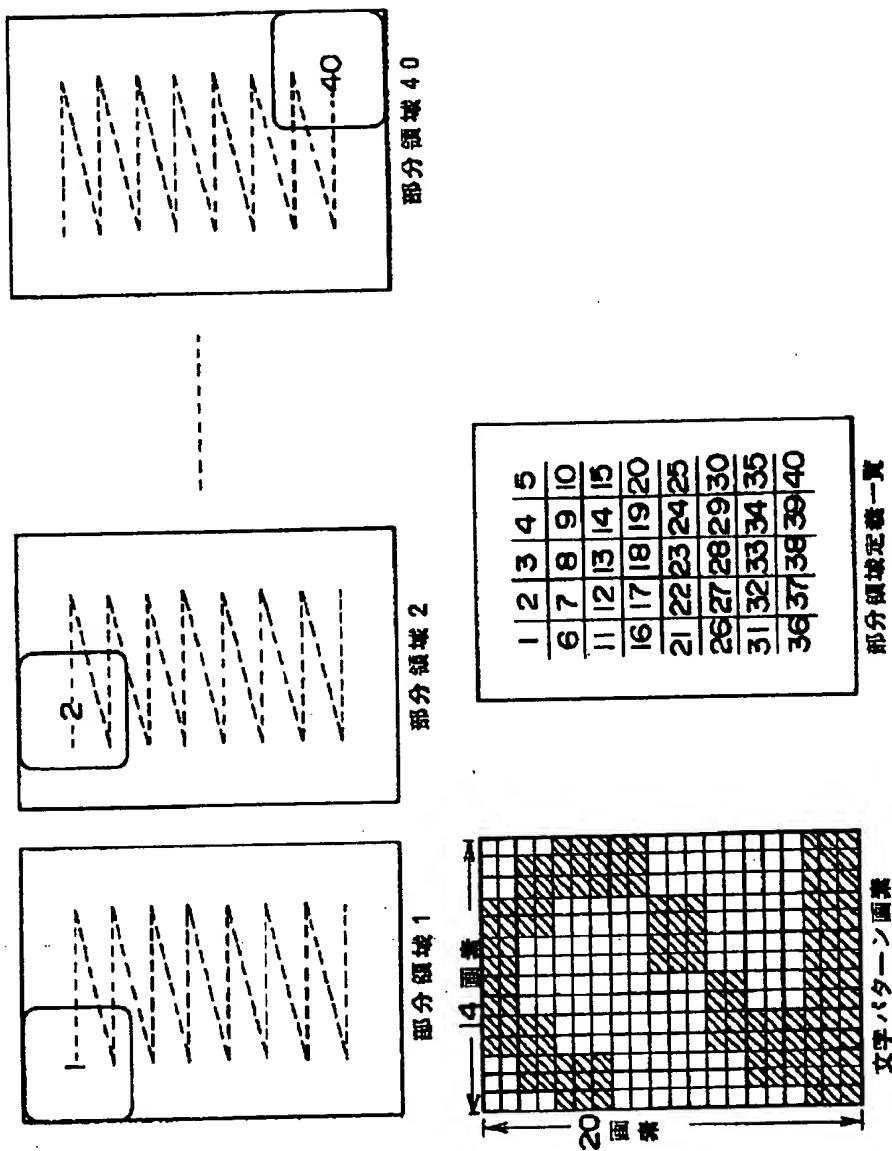
【図1】



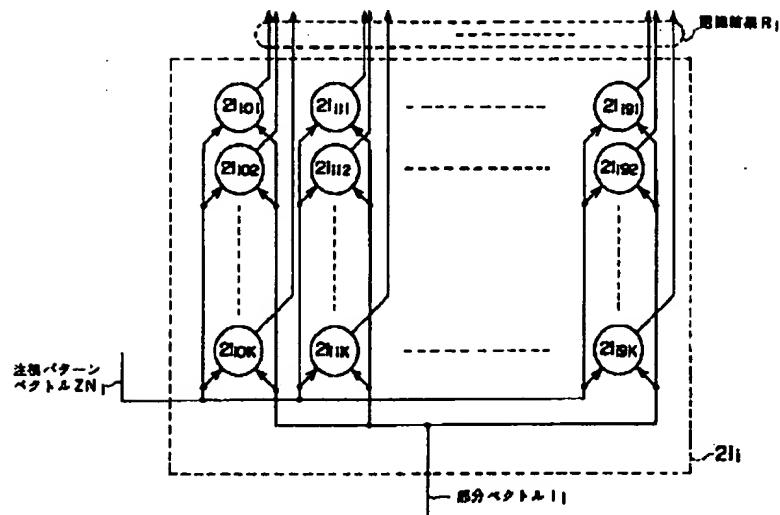
【図2】



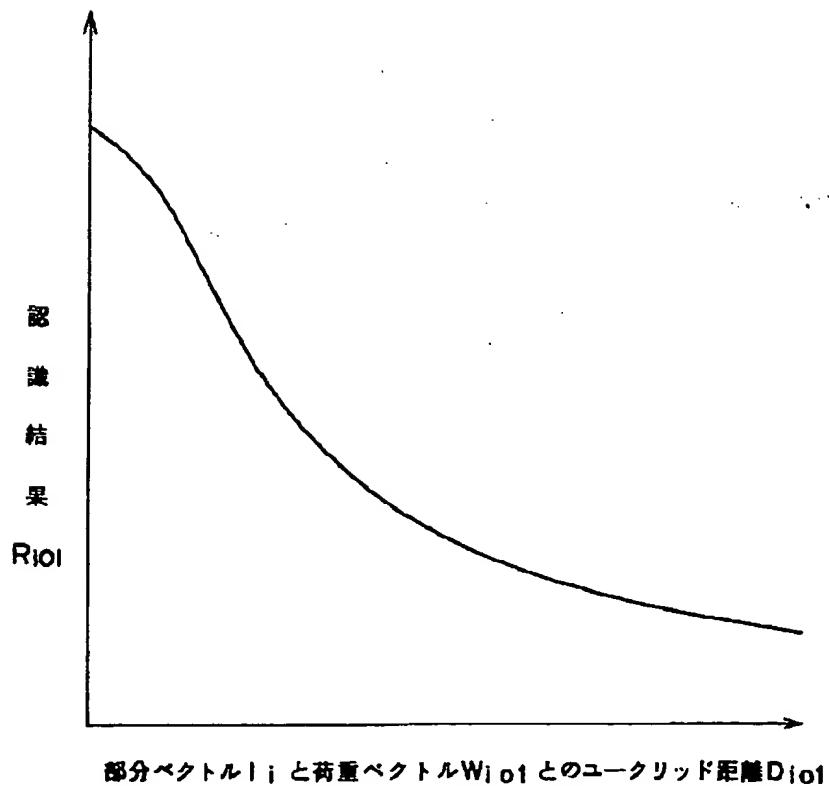
【図3】



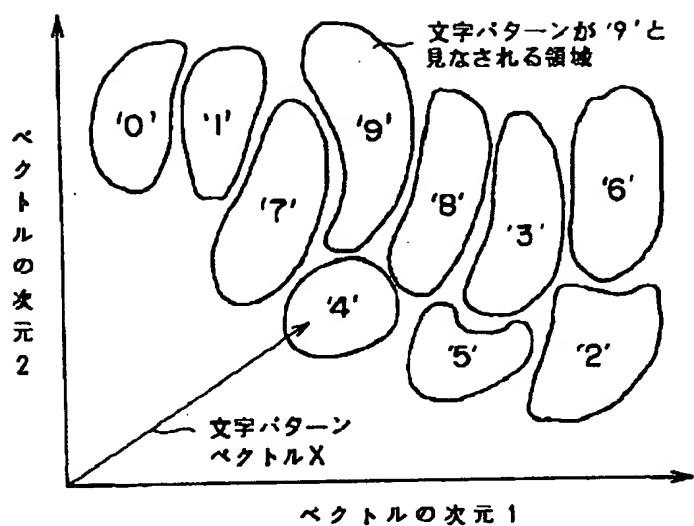
【図4】



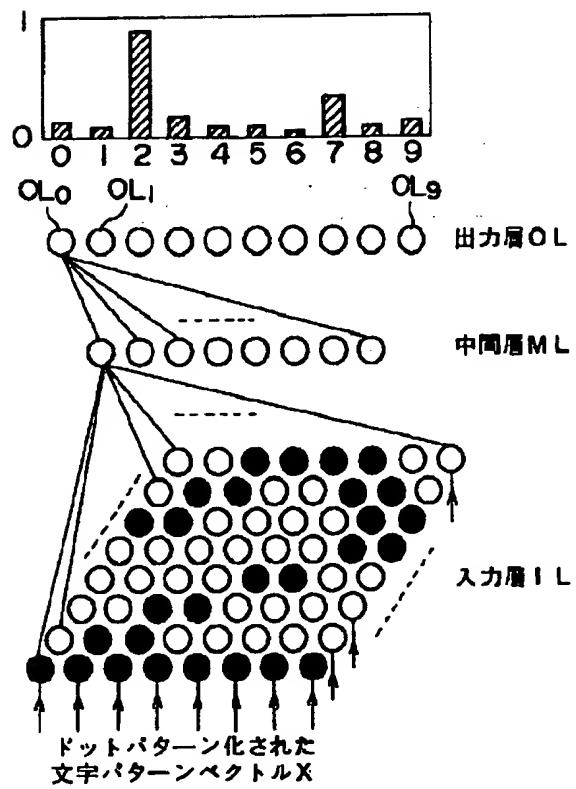
【図5】



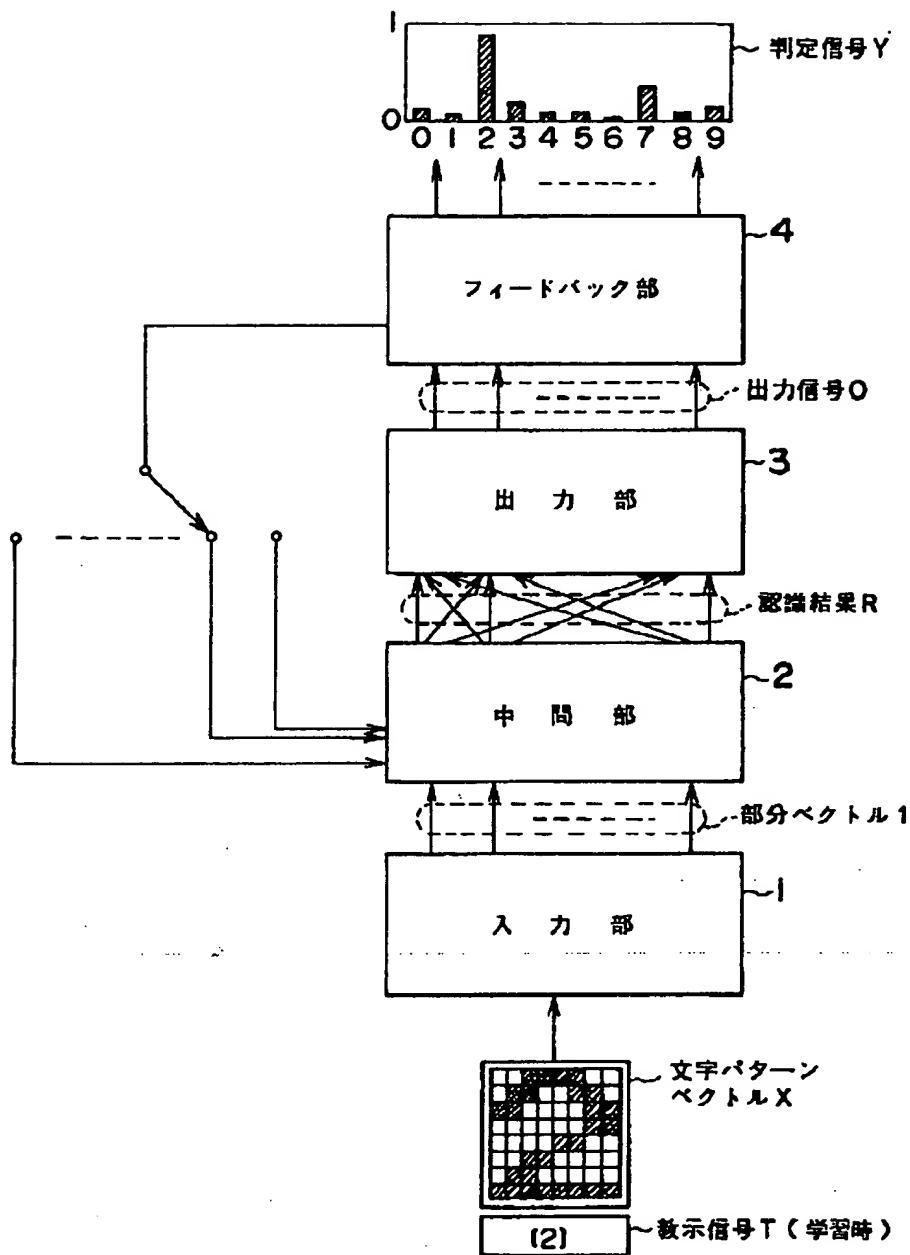
【図7】



【図8】



【図10】



フロントページの続き

(58)調査した分野(Int.Cl.5, DB名)

G06K 9/62 620

G06K 9/66

G06F 15/18 560

THIS PAGE BLANK (USPTO)